IDENTIFICATION OF GENDER AND FASTING GENES USING RAT LIVER AND KIDNEY BASELINE GENE EXPRESSION

Jeff Chou[1], Jie Liu[1], Susan Hester[4], Janice Lee[4], Karol Thompson[3], Mike Bass[2], Mike Boedigheimer[2], Jennifer Fostel[1], Fenglong Liu[5], Russ Wolfinger[6], Pierre Bushel[1], J. Chris Corton[4]

[1]NIEHS, Research Triangle Park, NC, [2]Amgen, Thousand Oaks, CA, [3]CDER, US FDA, Silver Spring, MD, [4]EPA, Research Triangle Park, NC, [5]Dana-Farber Cancer Institute, Boston, MA, [6]SAS, Cary, NC

Abstract

Microarray data from independent labs and studies can be compared to potentially identify toxicologically and biologically relevant genes. The Baseline Animal Database working group of HESI was formed to assess baseline gene expression from microarray data derived from control or vehicle-treated rats across multiple studies and sites. We determined whether this database could be used to identify genes clearly associated with a number of biological factors that were part of the database. We used Extracting Patterns and Identifying co-expressed Genes (EPIG) to identify gender-specific and fasting-specific genes from each institute in which the factor could be examined in isolation. Gender-specific genes were identified in liver and kidney from 5 different institutes using a total of 134 liver and 58 kidney transcript profiles. A total of 266 liver genes and 313 kidney genes were identified that exhibited similar behavior (both direction of change and degree of gender difference) in at least 2/4 and 2/3 institutes, respectively. Surprisingly, approximately 1/3 of the kidney genes were regulated in an opposite manner than in liver. The gender genes overlapped with well-characterized genes from earlier studies. Using in silico approaches, we identified a number of known response elements in the promoters of co-regulated genes including those involved in determining gender-specific gene expression. Fasting-specific genes were also identified in a similar manner and significantly overlapped with those that were known to be regulated by fasting in rodents. Our results demonstrate that biologically relevant genes that are associated with a number of variables common to toxicology studies can be identified using a well annotated data of microarray dataset from multiple sources. This abstract does not necessarily reflect EPA policy.

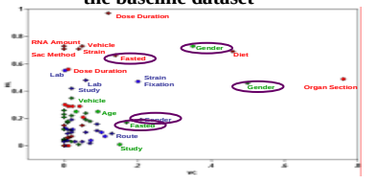
Introduction

A number of studies have shown the usefulness of databases of microarray information for chemical classification and mode of action analysis. These microarray databases are usually the result of parallel studies using similar if not identical conditions of animal treatment, RNA isolation and microarray analysis of RNA expression which facilitates the comparison of the effects of one chemical to another. However, most toxicogenomic studies are analyzed in relative isolation, i.e., in the absence of any direct comparison of the generated microarray data to archived microarray data. In the absence of any direct comparison the interpretation of the data can be biased and based on knowledge of the function and regulation of the altered genes as viewed by the interpreter.

There are a large number of archived microarray studies that have been submitted to sites such as GEO that could be potentially useful for the interpretation of a microarray experiment. However, the potentially large number of differences in study design and microarray platform could lead to misinterpretation of the transcript profiles.

In this study we determined the usefulness of microarray data generated in different labs and under different exposure scenarios to identify genes associated with gender or fasting. We used the dataset generated as part of the Baseline Assessment Database Subcommittee of the Health and Environmental Sciences Institute (HESI) Technical Committee on the Application of Genomics in Mechanism Based Risk Assessment (described in greater detail in the SOT poster #1215 Thompson et al.). Voluntary contributions of microarray data from the control arms of toxicogenomic studies for liver and kidney were requested from HESI member institutions including information on subject characteristics and husbandry, methods used for euthanasia, specimen preparation and preservation protocols, RNA preparation and labeling, and microarray hybridization. This poster describes the use of this dataset to expand our knowledge of genes that are associated with gender differences in the liver and kidney and the effect of fasting in the liver.

Gender and fasting contribute to variability in the baseline dataset



Major sources of variability according to two different multivariate statistics. HL = Hotelling-Lawley trace statistic (scaled relative to maximal model value), VC = Variance Components (computed based on principal components analysis. Affymetrix GeneChip type: RAE230_2 (●); RAE230A (●); RGU34A (●). Tissue: Kidney (●); Liver (+). The genes that determine the differences related to gender and fasting are analyzed in this study.

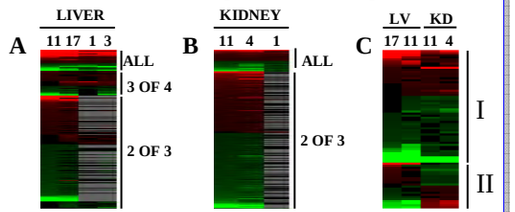
Source of data used to identify gender-selective genes

Institution	Study code	Gender	Tissue	Number of animals	Strain	Chip type	Fasted?
1	B	M	Liver	10	Wistar	U34A	No
1	C	F	Liver	4	Wistar	U34A	No
1	C	F	Kidney	4	Wistar	U34A	Yes
1	C	M	Liver	4	Wistar	U34A	No
1	C	M	Kidney	5	Wistar	U34A	Yes
3	A	M	Liver	3	Wistar	U34A	No
3	C	F	Liver	6	Wistar	U34A	No
3	C	M	Liver	6	Wistar	U34A	No
3	D	M	Liver	4	Wistar	U34A	No
4	B	F	Kidney	9	SD	RAE230_2	No
4	B	M	Kidney	10	SD	RAE230_2	No
11	A	F	Liver and Kidney	5	SD	230A	No
11	A	M	Liver and Kidney	5	SD	230A	No
11	B	F	Liver and Kidney	5	SD	230A	No
11	B	M	Liver and Kidney	5	SD	230A	No
11	C	F	Liver and Kidney	5	SD	230A	No
11	C	M	Liver and Kidney	5	SD	230A	No
17	A	F	Liver	15	SD	230A	No
17	D	F	Liver	4	SD	230A	No
17	E	F	Liver	14	SD	230A	No
17	F	M	Liver	16	SD	230A	No
17	G	M	Liver	18	SD	230A	No

Summary of analysis and results: gender genes

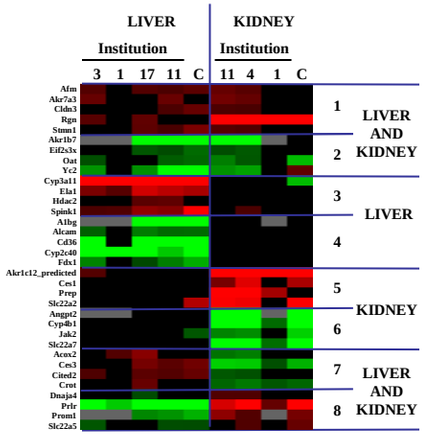
•Used the transcript profiles from a total of 14 studies carried out at 5 institutions (livers: 76 males and 58 females; kidneys: 30 male and 28 female rats.
•Analyzed using EPIG, a profile-based analysis method for Extracting microarray gene expression Patterns and Identifying co-expressed Genes (Zhou et al., 2006).
•A total of 854 or 863 genes were significantly different in one or more institutes in liver or kidney, respectively.
•Out of these ~31 or ~35% of the genes exhibited significant differences in more than one institute in liver or kidney, respectively.
•Overall, the genes showed excellent concordance between institutes in terms of gender specificity as well as magnitude of the difference between genders.

Identification of gender-selective genes



(A and B) Red indicates higher expression in males vs. females, green indicates higher expression in females vs. males, black indicates no significant difference and grey indicates no data due to lower genome coverage in the U34A chips. Only those genes that were altered in two or more of the three (kidney) or four (liver) institutions are shown.
(C) Two classes of gender-predominant genes. Genes identified in the liver or kidney were compared as described using only those genes identified from institutions that used the 230A or 230_2 chips and were altered in three or four of the four institutions examined.

Identification of different classes of gender-selective genes



Genes that exhibited significantly altered expression between males and females in liver or kidney were divided into 8 different groups based on expression behavior. Groups 1, 2, 7 and 8 exhibited differential expression in the liver and kidney whereas groups 3 and 4 or 5 and 6 exhibited liver-predominant or kidney-predominant differential expression, respectively. The expression of four or five genes from each of the groups was examined by RT-PCR (columns "C") in the livers and kidneys from control male and female F344 rats from an independent study.

Source of data used to identify fasting-specific genes

Institution	Study code	Gender	Strain	Number of animals	Chip type	Fasted?
2	A	M	Wistar	12	230A	No
2	D	M	Wistar	12	230A	No
2	E	M	Wistar	12	230A	No
7	A	M	SD	5	230A	No
8	B	F	Fisher	11	230A	Yes
11	A	F	SD	5	230A	No
11	A	M	SD	5	230A	No
11	B	F	SD	5	230A	No
11	B	M	SD	5	230A	No
17	C	F	SD	5	230A	Yes
17	D	F	SD	4	230A	Yes
17	F	M	SD	16	230A	Yes
17	G	M	SD	18	230A	Yes

Summary of analysis and results: fasting genes

•Used the transcript profiles from a total of 11 studies carried out at 5 institutions (61 animals fed ad libitum (AL) and 54 animals fasted).

•Genes associated with fasting were identified by EPIG and t-test ($p<0.05$).

•A total of 190 or 311 genes were significantly different using either EPIG or t-test, respectively.

Identification of fasting-specific genes

Genes that exhibited altered expression between ad libitum (AL) or fasted animals were identified using either the EPIG (A) or t-test (B). Genes and individual animals were clustered using two-dimensional clustering.

Gene Ontology analysis

Liver gender genes

Processes Name	Gene Abbreviations
Xenobiotic metabolism/ response to xenobiotic stimulus	AHR, CYP2C9, CYP3A4, GSTM1, CYP3A43, CYP1A7, CYP3A3
Cellular lipid metabolism/ lipid metabolism	Acyl-CoA synthetase, HSD17B2, AKR1C3, HSD11B1, CD36, FDX1, CYP3A4, Prolactin receptor, SULT1E1, CES1, PTGS structural, GFAM, SULT2A1, PI3K class II, ACSL5, HAOX2, STAR, PPP2R1B, NUDT4
Steroid metabolism	HSD17B2, AKR1C3, HSD11B1, FDX1, Prolactin receptor, SULT1E1, SULT2A1, STAR
Hormone metabolism	HSD17B2, AKR1C3, HSD11B1, SULT1E1, PACE4, STAR
Carboxylic acid and fatty acid metabolism	Acyl-CoA synthetase, AKR1C3, CD36, SLC7A7, CES1, PTGS structural, GFAM, SLC7A5, CSAD, ACSL5, HAOX2, PPP2R1B
Response to chemical stimulus	AHR, Carbonic anhydrase III, CYP2C9, CYP3A4,GSTM1, CYP3A43, CES1, CYP1A7, CYP3A3,PPZA structural, ENPP2, PPP2R1B

Kidney gender genes

Processes Name	Gene Names
lipid metabolism	PLD1, Angiotensinogen,Myosin Va,P-glycoprotein,HSD17B2,HSD11B1, ADHC, Alpha-synuclein,LPL, HPGD,IDI1,FDFT1, ACADM,Prolactinreceptor,CES1,PGHD,RBP1,Non-A beta component of AD amyloid, Alpha-synuclein filaments, AZGP1,Retinol-binding protein,3.TB4DH,CROT, ACO32, AAAD,SULT2B1,
aromatic compound metabolism	Occludin,FTCD,Alpha-synuclein,CYP1A1,HPD,Non-A beta component of AD amyloid,Alpha-synuclein filaments, GAMT,FAAA
carboxylic/organic acid metabolism	Angiotensinogen,Myosin Va,ACSL1,FTCD,LPL,HPGD,ACADM,CES1,ASNS,PGHD,GPD2,HPD,OAT,1.TB4DH,CROT,CSAD, ACO32,GAMT,FAAA
vitamin transport	Cubilin,SLC19A1,VDB,SMVT
response to drug	"MDR1,P-glycoprotein, Alpha-synuclein,Non-A beta component of AD amyloid,Alpha-synuclein filaments,GSTA1"
positive regulation of neurotransmitter secretion	AAlpha-synuclein,Non-A beta component of AD amyloid,Alpha-synuclein filaments
fatty acid metabolism	Angiotensinogen,Myosin Va,LPL,HPGD,ACADM,CES 1,PGHD,1.TB4DH,CROT,ACO32,
amino acid derivative biosynthesis	Alpha-synuclein,Non-A beta component of AD amyloid,Alpha-synuclein filaments,2.SAD,G AMT

Fasting genes

Processes Name	Gene Names
lipid metabolism	CYP222,PLA2,SCD,ACLY,.HMDH,HMGCS1,SREBP1,Acyl-CoA synthetase,MIF,CYP4F2,P-glycoprotein,MDR1,HSD17B2,Fatty acid-binding protein,CYP17, CPT-1A,Acetyl-CoA acyltransferase,IDI1,FDFT1,ERG1,LSS,DHCR7,FDPS,CYP51A1,SC4MOL,AFP,CRAT,B-FABP,CPT ILS14 protein,CYP4A11,.HADHB,ACSL3,P.AHX,E-FABP,NSDHL,.GPAM,ACAT2,PLA2G12,DC.LAFP-cyclopeptide,DHC24,PTE2,CYP4F11,CTE1,FXR,PMVK,HPLC2
lipid biosynthesis	PLA2,SCD,.HMDH,SREBP1,Acyl-CoA synthetase,MIF,HSD17B2,Fatty acid-binding protein,CYP17,.IDI1,FDFT1,ERG1,LDHCR7,FDPS,CYP51A1,SC4MOL,AFP,NSDHL,DHC24
steroid metabolism/ biosynthesis	HMDH,HSD17B2,CYP17,.IDI1,FDFT1,ERG1,LSS,DHCR7,FDPS,CYP51A1,SC4MOL,AFP,NSDHL,AFP-cyclopeptide,.DHC24,FXR
alcohol metabolism	HMDH,IBP1,6PGD,Pyruvate kinase,Fatty acid-binding protein,IDI1,FDFT1,ERG1,.DHCR7,FDPS,CYP51A1,.DOC,ADH4,.E-FABP,NSDHL,.IBP,DHC24
fatty acid oxidation	Fatty acid-binding protein,CPT-1A,Acetyl-CoA acyltransferase,CPT II,HADHB,PAHX,DCI,HPLC2

Identification of transcription factor binding sites in the promoters of gender genes

Three classes of transcription factor binding sites were identified using Genomatix. Binding sites for the following transcription factors were identified:
•STAT1, STAT3, STAT5, STAT6: some of these transcription factors control gender selective gene expression through growth hormone-JAK pathways.
•LXR-RXR, FXR-RXR, VDR-RXR: heterodimeric nuclear receptors that respond to oxysterols, bile acids and vitamin D, respectively.
•CAAT Displacement Protein (CDP) family members including Cut and Cux: homeodomain-containing proteins that repress the transcriptional activities of HNF-1α, C/EBPα and Rb/p107.

Conclusions

• We identified a comprehensive set of gender-specific genes in the liver which overlap with those previously characterized

•We identified gender-specific genes in the kidney. Genes with this pattern of expression have not been well documented in the literature.

•A set of gender-specific genes that exhibit opposite gender specificity in the liver and kidney were identified.

•A large set of fasting genes were identified.

•Meta analysis of a large dataset from different studies/labs can be used to identify biologically relevant genes.

The views in this poster do not necessarily reflect EPA policy.